Kernel density estimation on Riemannian Manifolds: Asymptotic Results.*

Guillermo Henry
Departamento de Matemática, FCEyN,
Universidad de Buenos Aires and CONICET, Argentina.
and
Daniela Rodriguez
Departamento de Matemática and Instituto de Cálculo, FCEyN,
Universidad de Buenos Aires and CONICET, Argentina

Abstract

The paper concerns the strong uniform consistency and the asymptotic distribution of the kernel density estimator of random objects on a Riemannian manifolds, proposed by Pelletier (2005).

Key words and phrases: statistics, geometry, Nonparametric estimation, Riemannian Manifolds.

1 Introduction

In recent years with the objective to explore the nature of complex nonlinear phenomena, the field of nonparametric inference has increased attention. The idea of nonparametric inference is to leave the data to show the structure lying beyond them, instead of imposing one. Kernel density estimation is a well-known method for estimating the probability density function of a random sample. However, in many applications, the variables $\mathbf{x}$ take values on a Riemannian manifold more than on $I\!R^d$ and this structure of the variables needs to be taken into account when considering neighborhoods around a fixed point $\mathbf{x}$. Some examples could be found in image analysis, astronomy, geology and other fields, they include distributions on spheres, orthogonal groups, Lie groups. Research on the statistical analysis of such objects was studied by [9], [4] and more recently [12] and [8]. Nonparametric kernel methods for estimating densities of spherical data have been studied by [7] and [1].

*This research was partially supported by Grants X-094 from the Universidad de Buenos Aires, PID 5505 from CONICET and PAV 120 and PICT 21407 from ANPCYT, Argentina.

Pelletier([10]) proposed a family of nonparametric estimators for the density function based on kernel weight when the variables are random object valued in a Riemannian manifolds. More precisely, let (M, g) be complete Riemannian manifolds and let us consider $\mathbf{x}_1, \ldots, \mathbf{x}_n$ independent and identically distributed random object on M with density function $f(p)$. The Pelletier's idea was to consider an analogue of a kernel on (M, g) by using a positive function of the geodesic distance on M, which is then normalized by the volume density function of (M, g) to take into account for curvature. These estimators are an average of the weight depending on the distance between $\mathbf{x}_i$ and p. The Pelletier's estimators is consistent with the kernel density estimators in the Euclidean case. Pelletier ([10]) studied L^2 convergence rates, under regularity conditions. The object of this note is to complement Pelletier's results with classical properties such as strong uniform consistency and asymptotic distribution.

This paper is organized as follows. Section 2 contains a brief summary of the Pelletier's proposal. Uniform consistency of the estimators is derived in Section 3, while in Section 4 the asymptotic distribution is obtained under regular assumptions on the bandwidth sequence. Proofs are given in the Appendix.

2 Pelletier's density estimator

2.1 Preliminaries

Let (M, g) be a $d-$dimensional oriented Riemannian manifold without boundary. Denote by d_g the distance induced by g and by $inj_g M = \inf_{p \in M} \sup\{s \in I\!R > 0 : B_s(p) \text{ is a normal ball}\}$ the injectivity radius of (M, g).

Througout this paper, we will assume that (M, g) is complete, i.e. (M, d_g) is a complete metric space, and that $inj_g M$ is strictly positive. Some examples of Riemannian manifolds with positive injectivity radius are $I\!R^d$ with g the canonical metric ($inj_g I\!R^d = \infty$), and the $d-$dimensional sphere S^d with the metric induced by the canonical metric of $I\!R^d$ ($inj_g S^d = \pi$). It is also well known that compact Riemannian manifolds have positive injectivity radius. Moreover, complete and simply connected Riemannian manifolds with non positive sectional curvature, have also this property. Some standard results on differential geometry can be seen for instance in [2], [3], [5] and [6].

From now on, we will denote by $B_s(p)$ the normal ball in (M, g) centered at p with radius s. Then, $B_s(0_p) = exp_p^{-1}(B_s(p))$ is an open neighborhood of 0_p in T_pM, the tangent space of M at p, and so it has a natural structure of differential manifold. We are going to consider the Riemannian metrics g' and g'' in $B_s(0_p)$, where $g' = exp_p^*(g)$ is the pullback of g by the exponential map and g'' is the canonical metric induced by g_p in $B_s(0_p)$. Let $w \in B_s(0_p)$, for any chart $(\bar{U}, \bar{\psi})$ of $B_s(0_p)$ such that $w \in \bar{U}$, the volumes of the parallelepiped spanned by $\{(\partial/\partial\bar{\psi}_1|_w), \ldots, (\partial/\partial\bar{\psi}_d|_w)\}$ with respect to the metrics g' and g'' are given by $|\det g'((\partial/\partial\bar{\psi}_i|_w), (\partial/\partial\bar{\psi}_j|_w))|^{1/2}$ and $|\det g''((\partial/\partial\bar{\psi}_i|_w), (\partial/\partial\bar{\psi}_j|_w))|^{1/2}$ respectively. The quotien between these two volumes is independent of the selected chart. So, given $q \in B_s(p)$, if $w = exp_p^{-1}(q) \in B_s(0_p)$ we can define the volume density function,

$\theta_p(q)$, on (M, g) as

$$\theta_p(q) = \frac{|\det g'\left((\partial/\partial\bar{\psi}_i|_w), (\partial/\partial\bar{\psi}_j|_w)\right)|^{1/2}}{|\det g''\left((\partial/\partial\bar{\psi}_i|_w), (\partial/\partial\bar{\psi}_j|_w)\right)|^{1/2}}$$

for any chart $(\bar{U}, \bar{\psi})$ of $B_s(0_p)$ that contains $w = exp_p^{-1}(q)$. For instance, if we consider the exponential chart (U, ψ) of (M, g) induced by an orthonormal basis $\{v_1, \ldots, v_d\}$ of T_pM and U a normal neighborhood of p then

$$\theta_p(q) = \left|\det g_q\left(\frac{\partial}{\partial\psi_i}\Big|_q, \frac{\partial}{\partial\psi_i}\Big|_q\right)\right|^{\frac{1}{2}},$$

where $\frac{\partial}{\partial\psi_i}|_q = D_{\alpha_i(0)}exp_p(\alpha_i'(0))$ with $\alpha_i(t) = exp_p^{-1}(q) + tv_i$ for $q \in U$. Note that the volume density function $\theta_p(q)$ is not defined for all p and q in M, but only for those points such that $d_g(p, q) < inj_gM$. It is worth noticing that, when M is $\mathbb{R}^d$ with the canonical metric, then $\theta_p(q) = 1$ for all $p, q \in \mathbb{R}^d$. See also, [3] and [12] for a discussion on the volume density function.

2.2 The estimator

Consider a probability distribution with a density f on a $d-$dimensional Riemannian manifold (M, g). Let $\mathbf{x}_1, \cdots, \mathbf{x}_n$ be i.i.d random object takes values on M with density f. In order to estimate f using observations $\mathbf{x}_1, \ldots, \mathbf{x}_n$ Pelletier ([1]) proposed the following kernel estimate:

$$f_n(p) = \frac{1}{nh_n^d}\sum_{i=1}^{n}\frac{1}{\theta_p(\mathbf{x}_i)}K\left(\frac{d_g(p, \mathbf{x}_i)}{h_n}\right) \tag{1}$$

where $K : \mathbb{R} \to \mathbb{R}$ is a non-negative function, $\theta_p(q)$ denotes the volume density function on (M, g) and the bandwidth h_n is a sequence of real positive numbers such that $\lim_{n\to\infty} h_n = 0$ and $h_n < inj_gM$, for all n. This last requirement on the bandwidth guarantees that (1) is defined for all $p \in M$ (see, [11]).

3 Consistency

Let U an open set of M, we denote by $C^k(U)$ the set of k times continuously differentiable functions from U to $\mathbb{R}$. As in [10], we assume that the image measure of P by $\mathbf{x}$ is absolutely continuous with respect to the Riemannian volume measure ν_g, and we denote by f its density on M with respect to ν_g. In this section we will consider the following set of assumptions to derive the strong consistency results of the estimate $f_n(p)$ definied by Pelletier in [10].

$H1$. Let M_0 be a compact set on M such that:

i) f is a bounded function such that $\inf_{p\in M_0} f(p) = A > 0$.

ii) $\inf_{p,q\in M_0} \theta_p(q) = B > 0$.

- $H2$. For any open set U_0 of M_0 such that $M_0 \subset U_0$, $f \in C^2(U_0)$.
- $H3$. The sequence h_n is such that $h_n \to 0$, $nh_n^d \to \infty$ and $\frac{nh_n^d}{\log n} \to \infty$ as $n \to \infty$.
- $H4$. $K : I\!R \to I\!R$ is a bounded nonnegative Lipschitz function of order one, with compact support $[0, 1]$ satisfying: $\int_{I\!R^d} K(\|\mathbf{u}\|)d\mathbf{u} = 1$, $\int_{I\!R^d} \mathbf{u}K(\|\mathbf{u}\|)d\mathbf{u} = \mathbf{0}$ and $0 < \int_{I\!R^d} \|\mathbf{u}\|^2 K(\|\mathbf{u}\|)d\mathbf{u} < \infty$.

Remark 3.1. The fact that $\theta_p(p) = 1$ for all $p \in M$ guarantees that $H3$ ii) holds. Assumption $H4$ is a standard assumption when dealing kernel estimators.

Remark 3.2. Using the Theorem 3.2 in [10] and the compactness of M_0 we have that

$$\sup_{p \in M_0} |E(f_n(p)) - f(p)| \leq c\, h_n^2 \int \|\mathbf{u}\|^2 K(\|\mathbf{u}\|)d\mathbf{u}.$$

Then, in order to obtain strong uniform consistency it suffices to show that

$$\sup_{p \in M_0} |f_n(p) - E(f_n(p))| \xrightarrow{a.s.} 0.$$

Theorem 3.3. Assume that $H1$ to $H4$ holds, then we have that

$$\sup_{p \in M_0} |f_n(p) - E(f_n(p))| \xrightarrow{a.s.} 0.$$

4 Asymptotic normality

Denote by $V \subset M$ an open neighborhood of p. With the objective to derive the asymptotic distribution of the estimator of f, we will consider two assumptions more.

- $H5$. $f(p) > 0$, $f \in C^2(V)$ and the second derivative is bounded.
- $H6$. The sequence h is such that $h_n \to 0$, $nh_n^d \to \infty$ as $n \to \infty$ and there exists $0 \leq \beta < \infty$ such that $\sqrt{nh_n^{p+4}} \to \beta$ as $n \to \infty$.

In the following we will denote by $\mathcal{V}_s$ the ball in $I\!R^d$ centered at the origin and of radius s.

Theorem 4.1. Assume $H4$ to $H6$. Then we have that

$$\sqrt{nh_n^d}(f_n(p) - f(p)) \xrightarrow{\mathcal{D}} \mathcal{N}(b(p), V(p))$$

with

$$b(p) = \frac{\beta}{2} \int_{\mathcal{V}_1} K(\|\mathbf{u}\|)u_1^2 \, d\mathbf{u} \sum_{i=1}^{d} \frac{\partial f \circ \psi^{-1}}{\partial u_i \partial u_i}|_{u=0}$$

and

$$V(p) = f(p)\int_{\mathcal{V}_1} K^2(\|\mathbf{u}\|)d\mathbf{u}$$

where $\mathbf{u} = (u_1, \ldots, u_d)$ and $(B_{h_n}(p), \psi)$ some exponential chart induced by an orthonormal basis of T_pM.

Remark 4.2. Note that the Pelletier's estimator converges at the same rate as the Euclidean kernel estimator.

Appendix

From now on, we will denote by $d\nu_g$ the usual volume element induced by g and the orientation of M.

Proof of Theorem 3.3: Let us begin by fixing some notation. Given $p \in M$, denote

$$V_j(p) = \frac{1}{\theta_p(\mathbf{x}_j)} K\left(\frac{d_g(p, \mathbf{x}_j)}{h_n}\right) - E\left(\frac{1}{\theta_p(\mathbf{x}_j)} K\left(\frac{d_g(p, \mathbf{x}_j)}{h_n}\right)\right),$$

let $S_n(p) = \sum_{j=1}^n V_j$. The fact that $E(V_j) = 0$, the kernel K is bounded and the volume density function satisfies $\theta_p(q) \geq B > 0$ for all $p, q \in \mathcal{M}_0$, we have that $|V_j(p)| < A_1$. Then, Bernstein's inequality implies that, for $n > n_0$ and for some positive constants α, we have

$$\sup_{p \in \mathcal{M}_0} P\left(\frac{1}{nh_n^p}|S_n(p)| > \varepsilon\right) \leq 2\exp\left(-nh_n^d\alpha\right). \tag{2}$$

On the other hand, since M_0 is a compact set, we can consider a finite collection of balls $(B_i = B_{h_n^\gamma}(p_i))$ centeres at points $p_i \in M_0$ with radius h_n^γ with $\gamma > 2 + d$, such that $M_0 \subset \cup_{i=1}^l B_i$. Then, $l = O(h_n^{-\gamma})$ and

$$\sup_{p \in M_0} |S_n(p)| \leq \max_{1 \leq j \leq l} \sup_{p \in B_j} |S_n(p) - S_n(p_j)| + \max_{1 \leq j \leq l} S_n(p_j)|. \tag{3}$$

Using that K is Lipschitz function with Lipschitz constant $\|K\|_L$, straightforward calculation lead to

$$\frac{1}{nh_n^d}|S_n(p) - S_n(p_j)| < 2\|K\|)L\frac{1}{nh_n^d}nh_n^{\gamma-1} = Ch_n^{\gamma-(d+1)}$$

for all $p \in B_j$, which entails that for n large enough, let us say, for $n > n_1$, we have

$$\max_{1 \leq j \leq l} \sup_{p \in B_j} \frac{1}{nh_n^d}|S_n(p) - S_n(p_j)| < \varepsilon \tag{4}$$

Finally, (2), (3) and (4) implies that, for $n > \max\{n_0, n_1\}$

$$P\left(\sup_{p \in M_0} \frac{1}{nh_n^d}|S_n(p)| > 2\varepsilon\right) \leq P\left(\max_{1 \leq j \leq l} |S_n(p_j)| > \varepsilon\right) \leq lh_n^{-\gamma}n^{\alpha\delta_n}$$

with $\delta_n = \frac{nh_n^d}{\log n}$. Taking $\gamma = d+3$ and using, that from $H3$, $nh_n \to \infty$ and $\delta_n \to \infty$, we have that for $n > n_2$, $(nh_n)^{-\gamma} < 1$ and $\gamma - \delta_n \alpha < 2$. Hence, for $n \geq \max\{n_0, n_1, n_2\}$ and some constant C', we get

$$P\left(\sup_{p \in M_0} \frac{1}{nh_n^d} |S_n(p)| > 2\varepsilon\right) \leq C' n^{-2}$$

which shows that $\sum_{n=1}^{\infty} P\left(\sup_{p \in M_0} \frac{1}{nh_n^d} |S_n(p)| > 2\varepsilon\right) < \infty$, concluding the proof. □

Proof of Theorem 4.1: Let $S_n(p) = \sum_{j=1}^n V_j(p)$ like in the previous theorem, with $V_j(p) = \frac{1}{\theta_p(\mathbf{x}_j)} K\left(\frac{d_g(p,\mathbf{x}_j)}{h_n}\right) - E\left(\frac{1}{\theta_p(\mathbf{x}_j)} K\left(\frac{d_g(p,\mathbf{x}_j)}{h_n}\right)\right)$. Firstly we note that if we take a Taylor expansion of f arround p at order two we get

$$\sqrt{nh_n^d}(E(f_n(p)) - f(p)) = \frac{\sqrt{nh_n^d}}{2} \sum_{i,j=1}^{d} \left[\frac{\partial f \circ \psi^{-1}}{\partial u_i u_j}|_{\mathbf{u}=\mathbf{0}} \int_{\mathcal{V}_1} K(\|\mathbf{u}\|) u_i u_j d\mathbf{u}\right] h_n^2 + \sqrt{nh_n^p}\, o(h_n^2)$$

where $\mathbf{u} = (u_1, \ldots, u_d)$ and $(B_h(p), \psi)$ some exponential chart induced by an orthonormal basis of T_pM. Then, the fact that $\int_{\mathcal{V}_1} K(\|\mathbf{u}\|) u_i u_j d\mathbf{u} = 0$ and $\int_{\mathcal{V}_1} K(\|\mathbf{u}\|) u_i^2 d\mathbf{u} = \int_{\mathcal{V}_1} K(\|\mathbf{u}\|) u_j^2 d\mathbf{u}$ if $i \neq j$ implies that

$$\sqrt{nh_n^d}(E(f_n(p)) - f(p)) \to \frac{\beta}{2} \sum_{i=1}^{d} \frac{\partial f \circ \psi^{-1}}{\partial u_i \partial u_i}|_{\mathbf{u}=\mathbf{0}} \int_{\mathcal{V}_1} K(\|\mathbf{u}\|) u_1^2 d\mathbf{u}$$

Therefore, it is enough to obtain the asymptotic behavior of $\frac{1}{\sqrt{nh_n^p}} S_n(p)$. In order to prove this, we will show that $V_j(p)$ satisfies the Linderberg Central Limit Theorem.

Pelletier ([11]) obtained that

$$\frac{1}{nh_n^d} \operatorname{var}\left(\sum_{j=1}^{n} V_j(p)\right) = f(p) \int_{\mathcal{V}_1} K^2(\|\mathbf{u}\|) d\mathbf{u} + o(1).$$

Finally, we note that

$$\begin{aligned} h^{-d} E(V_1^2(p) I(\mathbf{x}_1)) &\leq h_n^{-d} \int \frac{1}{\theta_p^2(q)} K^2\left(\frac{d_g(p,q)}{h_n}\right) I(q) f(q)\, d\nu_g(q) \\ &+ h_n^{-d} \left[E\left(\frac{1}{\theta_p(\mathbf{x}_1)} K\left(\frac{d_g(p,\mathbf{x}_1)}{h_n}\right)\right)\right]^2 \\ &= A_n + B_n \end{aligned}$$

where $I(q) = 1$ if $q \in \left\{q : |\frac{1}{\theta_p(q)} K\left(d_g(p,q)/h_n\right) - E\left(\frac{1}{\theta_p(\mathbf{x}_1)} K\left(d_g(p,\mathbf{x}_1)/h_n\right)\right)| > \sqrt{nh_n^d}\varepsilon\right\}$ and 0 in other case. Since

$$E\left(\frac{1}{\theta_p(\mathbf{x}_1)} K\left(\frac{d_g(p,\mathbf{x}_1)}{h_n}\right)\right) \leq h_n^d f(p) + C h_n^{2+d} \int_{\mathcal{V}_1} \|\mathbf{u}\|^2 K(\|\mathbf{u}\|) d\mathbf{u}$$

we have that $B_n \rightarrow 0$ as $n \rightarrow \infty$. Also, exists n_0 such that for all $n \geq n_0$ $|V_1(p)| \leq \frac{1}{\theta_p(\mathbf{x}_1)} K\left(\frac{d_g(p,\mathbf{x}_1)}{h_n}\right) + \frac{\varepsilon}{2}\sqrt{nh_n^d}$ then,

$$A_n \quad \leq \quad \int \frac{1}{h^d \theta_p^2(q)} K^2\left(\frac{d_g(p,q)}{h_n}\right) \widetilde{I}(q) f(q)\, d\nu_g(q)$$

where $\widetilde{I}(q) = 1$ if $q \in \left\{ q : \frac{1}{\theta_p(q)} |K(d_g(p,q)/h_n)| > \sqrt{nh_n^d}\varepsilon/2 \right\}$ and 0 in other case. Therefore, the fact that $\int \frac{1}{h_n^d \theta_p^2(q)} K^2\left(\frac{d_g(p,q)}{h_n}\right) f(q)\, d\nu_g(q) = f(p) \int_{\mathcal{V}_1} K^2(\|\mathbf{u}\|) d\mathbf{u} + o(1) < \infty$ implies that $A_n \rightarrow 0$ as $n \rightarrow \infty$ and $\sum_{i=1}^n E(\frac{1}{nh_n^d} V_j^2(p) I(\mathbf{x}_j)) \rightarrow 0$ as $n \rightarrow \infty$. Then, we conclude that $\sqrt{nh_n^d} S_n(p) \xrightarrow{\mathcal{D}} N(0, f(p) \int K^2(\|\mathbf{u}\|) d\mathbf{u})$. $\square$

References

[1] Bai, Z.D.; Rao, C. and Zhao, L. (1988). Kernel Estimators of Density Function of Directional Data. *J. Multivariate Anal.* **27,** 24-39.

[2] Berger, M.; Gauduchon, P. and Mazet, E. (1971). Le Spectre d' une variété Riemannienne. *Springer-Verlag.*

[3] Besse, A. (1978). Manifolds all of whose Geodesics are Closed. *Springer-Verlag.*

[4] Bhattacharya, R. and Patrangenaru, V. (2002). Nonparametric estimation of location and dispersion on Riemannian manifolds. *Journal of Statistical Planning and Inference.* **108,** 23-35.

[5] Boothby, W. M. (1975). *An introduction to differentiable manifolds and Riemannian geometry.* Academic Press, New York.

[6] Do Carmo, M. (1988). *Geometria Riemaniana.* Proyecto Euclides, IMPA. 2da edición.

[7] Hall, P. , Watson, G.S. and Cabrera, J. (1987). Kernel density estimation with spherical data. *Biometrika* **74**, 751762.

[8] Hendriks, H. and Landsman, Z. (2007). Asymptotic data analysis on manifolds. *Annals of Statistics,* **35, 1,** 109-131.

[9] Mardia, K. (1972). Statistics of Directional Data. *Academic Press, London.*

[10] Pelletier, B. (2005). Kernel Density Estimation on Riemannian Maniflods. *Statistics and Probability Letters*, **73, 3**, 297-304.

[11] Pelletier, B. (2006). Nonparametric regression estimation on closed Riemannian manifolds. *Journal of Nonparametric Statistics*, **18**, 57-67.

[12] Pennec, X. (2006). Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements. *Journal Math. Imaging Vis.,* **25,** 127-154.

Guillermo Henry
Departamento de Matemática, FCEyN, Universidad de Buenos Aires
Ciudad Universitaria, Pabellón I, Buenos Aires, C1428EHA, Argentina
e-mail address: ghenry@dm.uba.ar

Daniela Rodriguez
Instituto de Cálculo, FCEyN, Universidad de Buenos Aires
Ciudad Universitaria, Pabellón II, Buenos Aires, C1428EHA, Argentina
e-mail address: drodrig@dm.uba.ar